Suspicious Activity Detection Using Deep Learning

Uppala Sobhini[1], G. Lohith Sharma[2], K. Bhargavi[3], Garrepalli Shivani[4], C. Surekha[5]
[1,2,3,4] *UG Student, Computer Science in Artificial Intelligence & Machine Learning, Hyderabad Institute of Technology and management Gowdavelli Village, Medchal, Hyderabad, India*
[5]*Associate Professor, Hyderabad Institute of Technology and Management Gowdavelli Village, Medchal, Hyderabad, India*

***Abstract*—This project focuses on developing a machine learning model using deep learning techniques, specifically YOLO (You Only Look Once) and CNNs (Convolutional Neural Networks), to detect and recognize suspicious activities from surveillance images. The model is trained on the Dangerous Action Detection dataset, which includes classes such as weapons (knives, crowbars, guns) and suspicious human poses. The objective is to accurately identify potentially dangerous behaviors in real-time, improving safety in public and sensitive areas like airports, banks, and schools, where continuous monitoring is difficult. Intelligent video surveillance enables automatic distinction between normal and abnormal activities, triggering alerts to prevent crimes such as theft, vandalism, and terrorism. The project reviews advancements in suspicious activity detection over the past decade, addressing challenges in visual surveillance, including object detection, activity recognition, and foreground extraction. The results emphasize the effectiveness of deep learning models in enhancing crime prevention through automated surveillance systems. This solution contributes to public safety by enabling rapid, accurate detection of threats, reducing response times, and improving security measures.**

Keywords: Suspicious Activity Detection, Deep Learning, Visual Surveillance, YOLO, CNN, Crime Prevention, Automated Surveillance.

***Index Terms*—Deep Learning, YOLO, CNN, Real-Time Surveillance, Visual Surveillance, Object Detection, Human Activity Recognition, Abnormal Behavior Detection, Dangerous Action Detection Dataset, Automated Alert System, Feature Extraction, Crime Prevention, Intrusion Detection, Computer Vision, Smart Surveillance, Public Safety, Security Monitoring, Foreground Extraction, Video Analytics.**

I. INTRODUCTION

In today's rapidly urbanizing world, ensuring public safety has become a top priority. With increasing population density and the growing complexity of urban environments, maintaining secure public spaces presents significant challenges. Surveillance systems play a crucial role in safeguarding these areas by providing continuous monitoring and facilitating the early detection of potential threats. However, traditional surveillance systems heavily depend on human operators, making them prone to limitations such as fatigue, reduced attention spans, and human error. These factors can lead to delayed responses or missed critical events, diminishing the effectiveness of real-time threat management.

To overcome these limitations, this project explores the integration of advanced deep learning techniques, specifically YOLO (You Only Look Once) and CNNs (Convolutional Neural Networks), to automate and enhance suspicious activity detection in surveillance systems. YOLO is known for its real-time object detection capabilities, enabling the rapid identification of potential threats such as weapons or suspicious behavior. Meanwhile, CNNs excel in recognizing complex patterns, making them highly effective for detecting subtle anomalies in human behavior or identifying dangerous objects. This project highlights the transformative potential of artificial intelligence in public safety. The fusion of AI and surveillance not only enhances security but also creates a proactive, data-driven approach to crime prevention, ultimately contributing to more resilient and protected public spaces.

II. LITERATURE REVIEW

Intrusion Detection and Suspicious Activity Recognition

the effectiveness of intrusion detection and suspicious activity recognition has become a key area of focus in cybersecurity and visual surveillance. With the growing demand for real-time threat detection, researchers have increasingly explored machine learning and deep learning models to enhance surveillance accuracy and efficiency. This section reviews the relevant literature that forms the foundation for this project.

1. Survey of Intrusion Detection Systems: Techniques, Datasets, and Challenges
Source: Cybersecurity Journal

This study provides a comprehensive analysis of intrusion detection systems (IDS), emphasizing their role in identifying unusual patterns, which is crucial for detecting suspicious activities in network environments. The authors explore various IDS techniques, such as signature-based and anomaly-based detection methods. Signature-based IDS identify known threats by matching patterns to a database of previously identified attack signatures. In contrast, anomaly-based IDS detect deviations from normal system behavior, making them more effective in identifying new or unknown threats. The study also discusses the datasets and challenges associated with developing robust IDS, such as the need for diverse and comprehensive training data. The relevance of this research lies in its focus on anomaly detection, a principle that is directly applicable to visual surveillance systems, where deviations from typical behavior can indicate suspicious activities.

2. Anomaly-Based Intrusion Detection: Techniques, Systems, and Challenges
Source: ResearchGate
This paper explores anomaly-based detection methods and their application in identifying suspicious activities within computer networks. The study focuses on techniques such as statistical methods, machine learning algorithms, and clustering techniques, which are used to recognize deviations from expected behavior. These approaches are highly effective in real-world scenarios, where identifying deviations can help detect security threats. The paper highlights the importance of feature selection and extraction in improving the accuracy of anomaly detection systems. The insights from this research are valuable for visual surveillance applications, as recognizing unusual activity patterns in video footage is essential for detecting potential threats.

3. Deep Learning for Suspicious Activity Detection: A Survey
Source: ResearchGate
This study reviews how deep learning techniques are applied to detect suspicious activities, with a focus on video surveillance and cybersecurity. It highlights the advantages of models such as YOLO (You Only Look Once) and CNNs (Convolutional Neural Networks) for their speed and accuracy in processing large-scale visual data. YOLO is particularly effective for real-time object detection, making it ideal for identifying potential threats such as weapons or suspicious movements. CNNs, on the other hand, excel in recognizing intricate patterns, which enhances their ability to detect complex human behaviors and subtle anomalies. The paper underscores the growing role of deep learning in enhancing the accuracy and efficiency of visual surveillance systems.

III. METHODOLOGY

The primary goal of this project is to develop a deep learning-based system for detecting suspicious activities in real-time. Potential threats such as weapons (like knives and guns) and abnormal human behaviours (like hostility and thievery postures) are the primary focus of the system. A major problem is ensuring reliable detection while maintaining real-time speed, especially for applications used in high-security environments like public areas, banks, and airports. The project's dataset is sourced from Roboflow Universe's Dangerous Action Detection dataset, which includes tagged photos of potentially harmful objects and suspicious human behaviour. To enhance the model’s performance, the dataset undergoes multiple preprocessing steps.

2.1. Data preprocessing:
Data cleaning involves the elimination of incorrectly labeled and noisy data, while missing values are addressed through interpolation methods. Feature

engineering identifies significant suspicious behaviors, such as erratic movements and hidden objects.Normalization adjusts the pixel values of images to a range between 0 and 1, facilitating quicker convergence. Categorical encoding modifies object labels to make them suitable for deep learning applications. Ultimately, the dataset is divided into training (80%), validation (10%), and testing (10%) subsets to promote efficient learning and assessment.

2.2. Model Training and Evaluation:

The model was developed utilizing YOLO for real-time detection and CNNs for the extraction of features. Hyperparameter tuning was employed to enhance detection accuracy, and the optimal model was chosen based on metrics such as precision, recall, and F1-score. The evaluation process encompassed accuracy assessment, cross-validation, and testing on previously unseen data to guarantee reliability and generalization.

2.3. TEST METHODOLOGY:

The model underwent real-time surveillance testing in diverse environments, including low-light settings and densely populated areas, to guarantee precise detection of weapons, loitering activities, and unauthorized access. Functional testing validated the system's consistent classification capabilities across various object sizes and movements. Performance testing assessed the system's proficiency in managing multiple camera feeds while reducing latency on edge devices. Security testing confirmed the safety of data processing and the system's resilience against adversarial attacks. Lastly, usability testing ensured that alerts were clear and the system was responsive in high-security settings such as airports and financial institutions.

IV.ARCHITECTURE

The proposed method builds an automated framework that can quickly detect and classify suspicious activities using cutting-edge deep learning models. This architecture's careful design rapidly analyzes video inputs to provide precise object detection, accurate activity classification, and the generation of timely warnings to ensure prompt reactions. The system effortlessly integrates multiple components using a pipeline-based technique, each of which contributes to the overall objective of creating a reliable and scalable surveillance solution.

4.1 The System Architecture:

The distinct levels that comprise the architecture each have distinct functions. A more straightforward and effective monitoring system is produced when these layers work together to assess input data, extract useful features, classify actions, and issue the necessary alerts.Layer of Input:An input layer at the beginning of the system receives preprocessed image or video frames from security cameras. Preprocessing ensures that the unprocessed data is prepared for subsequent processing. It comprises decreasing noise, normalizing pixel values for consistency, and shrinking frames to a standard resolution in order to increase the clarity of visual data. The frames can be analyzed in batches or sequentially, depending on how the system is set up, ensuring continuous operation in real-time environments.Once the input has been produced, the feature extraction layer uses advanced models like Convolutional Neural Networks (CNNs) and YOLO (You Only Look Once) to identify objects and human activities inside each frame. CNNs extract spatial data like edges, textures, and other characteristics to detect objects and certain positions.patterns with remarkable accuracy. On the other hand, YOLO is employed because of its ability to detect several objects in a single frame. Its real-time processing capabilities ensures a balance between speed and accuracy for high-demand surveillance systems. Together, these models analyze the video frames to find relevant features for additional classification.

4.2 Classification Level:

The Classification Layer receives the features after processing and assigns labels to the recognized objects and activities. In order to distinguish between suspicious and typical activity, this layer makes use of sophisticated classification models that have been refined. For instance, the system can identify potentially harmful objects like weapons (like knives, guns, or crowbars) or unusual human behavior (such violence, vandalism, or trespassing). This layer reduces false positives to enhance detection.advantages and disadvantages.the system's dependability and reduces unnecessary distractions for security personnel.

4.3 Alert System:

The final component, the Alert System, activates if suspicious activity is discovered.Depending on the use case, the system can transmit alerts as messages to security personnel, interact with centralized monitoring systems, or start automated audio alarms. The alert system prioritizes critical situations, ensuring timely notice and enabling timely response. This function significantly enhances situational awareness and ensures real-time responsiveness in high-security scenarios.

4.4 Selecting a Model:

The effectiveness of the proposed system depends on selecting appropriate deep learning models that balance scalability, speed, and accuracy. Two basic models serve as the foundation for this architecture:

1. YOLO, or "You Only Look Once":

The greatest choice for real-time1 is YOLO. YOLO stands for "You Only Look Once."YOLO is the ideal choice for real-time object recognition due to its single-pass architecture, which processes images or video frames in a single forward pass. It is this design that makes YOLO so fast.and efficient. This system uses YOLO to recognize guns and other objects of interest.Because the system can accurately find and identify many items inside a single frame thanks to its bounding box predictions, it is highly suited for dynamic and complicated situations.

2. Convolutional neural networks, or CNNs:

CNNs are renowned for their capacity to identify and categorize features, especially when handling intricate patterns and background data in images. CNNs examine the frames of video in thisstructure to extract specific characteristics that enable abnormal human

V.RESULT ANALYSIS

The Suspicious Activity Detection System used modern deep learning algorithms like CNN and YOLO to identify suspicious activities and weapons with remarkable outcomes. With an overall accuracy of over 90%, the model accurately identified risks such as threatening movements, theft-related actions, and weapon identification. The detection process became more reliable as a result of the optimization of precision and recall scores, which significantly reduced false positives and negatives. The model is well-suited for real-world applications in environments such as airports, banks, and educational institutions, owing to its demonstrated resilience under a range of environmental factors, including overcrowded conditions and varying lighting scenarios. The system efficiently managed challenges like occlusions and different object orientations, enhancing reliability. These results demonstrate the ability for public safety and crime prevention, ensuring a scalable security solution in high-risk places.

VI. SCOPE FOR IMPROVEMENT

The project establishes a strong foundation for improving automated surveillance and the detection of suspicious activities, offering numerous prospects for future developments. Incorporating a wider range of environments, including offices, public transportation centres, and residential neighbourhoods in the dataset, will improve the model's ability to adapt to real-world situations. Additionally, the incorporation of several cameras can enhance detection accuracy by providing a comprehensive view of the monitored regions, facilitating coordinated tracking and the reconstruction of three-dimensional environments to improve spatial awareness. An automated alert system linked to law enforcement networks can enable rapid responses to detected threats, significantly improving public safety. These changes will increase the system's reliability and scalability, making it a critical tool for intelligent surveillance and crime prevention in high-risk zones.

VII. CONCLUSION

This project successfully demonstrates how deep learning can be used to enhance surveillance systems for the instantaneous detection of suspicious activity.By employing YOLO and CNN-based frameworks, the system effectively identifies weapons, aggressive postures, and a range of unusual behaviours with exceptional accuracy. The model's capability to analyze data in real time supports proactive crime prevention and improves security in high-risk environments.Although the existing system operates effectively in controlled environments, future improvements in dataset diversity, integration of multiple cameras, and real-time deployment at the edge will significantly boost its performance. This

research project provides the foundation for increasingly advanced and automated surveillance systems by highlighting the critical role of artificial intelligence in public safety.

REFERENCES

[1] X. Wang et al., "Unsupervised learning approach for abnormal event detection in surveillance video by revealing infrequent patterns." IEEE Xplore.
[2] J. A. Mohamed et al., "Spatial–temporal convolutional neural networks for anomaly detection and localization in crowded scenes." ScienceDirect.
[3] R. A. R. A. Kadir et al., "Survey of intrusion detection systems: techniques, datasets, and challenges." Springer Open.
[4] S. Y. Z. Nasir et al., "Anomaly-Based Intrusion Detection: Techniques, Systems, and Challenges." ResearchGate.
[5] L. M. S. Prasad et al., "Deep Learning for Suspicious Activity Detection: A Survey." ResearchGate.
[6] B. C. J. Tiwari et al., "A novel approach for suspicious activity detection with deep learning." Springer.
[7] P. J. G. Wang et al., "Human abnormal behavior detection using CNNs in crowded and uncrowded surveillance – A survey." ScienceDirect.
[8] X. Liu et al., "Reviewing approaches and techniques for detecting suspicious human behavior: A comprehensive survey." Elektrika Journal.